STABILITY OF COMPOSITE ESTIMATORS: EXPERIMENTS WITH HUNGARIAN LFS DATA*

ÖDÖN ÉLTETŐ – LÁSZLÓ MIHÁLYFFY

Since the establishment of the Central Statistical Office in 1867, the endeavour to keep pace with methodological developments of statistical bureaus and agencies leading the field has been a traditional feature of Hungarian official statistics. Application of survey methodology in social and agricultural statistics in Hungary since the fifties of this century as well as the results of those applications represented a standard which was acknowledged by the international community of official and survey statisticians. It is a commonplace that our period after the enormous changes in Mid-Eastern Europe and in the former Soviet Union in the early nineties unceasingly creates new challenges for official statisticians, especially in the so-called transition countries like Hungary. One way to respond to those challenges is to make no stop in improving and enhancing our methodological tools. The purpose of this paper is to take a step in this direction in the ranks of Hungarian official statistics.

Using composite estimators as a device to reduce the variance of direct sample estimates was introduced in the Current Population Survey (CPS) of the United States, in the beginning of the seventies.[1] The first version of those estimators labelled today as "the simple composite estimator" is basically the same which is considered below as a possible technique to be included in the processing of data of the Hungarian Labour Force Survey (LFS); the difference between estimators reflects that between sample designs. Since the initial "simple" stage, the composite estimator in the CPS has been the subject of many researches resulting in new and more efficient composite estimators.[2] Considering this development of the composite estimator and its possible introduction in the Hungarian LFS, it might seem somewhat odd to use the first generation of the method in this experiment rather than the last. The explanation of this approach is the

* The paper was presented at the IASS/IAOS Satellite Meeting on Longitudinal Studies. Jerusalem. August 27–31. 1997. The research was supported by the Hungarian Scientific Research Fund. No. T 017981.

[1] The Current Population Survey. Design and Methodology. US Bureau of the Census. Technical Papers 40.

[2] *Gurney, M. – Daly, J. F.*: A multivariate approach to estimation in periodic sample surveys. Proceedings of the Social Statistics Section. American Statistical Association. 1965. 242–257. p.; *Wolter, K. M.*: Composite estimation in finite population. *Journal of the American Statistical Association*. 1979. Vol. 74. 604–613. p.; *Breau, P. – Ernst, L. R.*: Alternative estimators to the current composite estimator. Proceedings of the Section on Survey Research Methods. American Statistical Association. 1983. 397–402. p.; *Cantwell, Patrick J.*: Variance formulae for composite estimation in rotation designs. Survey Methodology. Vol. 16. 153–163. p.

following. On the one hand, the simple composite estimator fits well in the current conditions of the Hungarian LFS; on the other hand, without certain modifications of the current practice of collecting and processing data of the LFS, the implementation of a sophisticated composite estimator would be meaningless, since the conditions of its efficiency are not fulfilled.

Some properties of the Hungarian Labour Force Survey

As was mentioned above, our purpose is to examine the conditions of possible introduction of a simple composite estimator in the Hungarian Labour Force Survey (LFS). To this end it is useful to have a brief review of the survey to see if it meets some basic conditions which are inevitable to define a composite estimator.

The LFS is a quarterly survey of households, which is based on a stratified probability sample of dwellings. The sample consists of a self-representing and a non-self-representing part, which were selected by a two-stage and a three-stage design, respectively. The dwellings or rather the addresses in the working sample are selected from the stock of addresses of 8272 enumeration districts (ED's) of the 1990 population census; this collection of ED's is called master sample and was designed to provide a sufficient number of addresses for the LFS in the period January 1992 – December 2001. The current practice of data collection is as follows:

– data are collected monthly, in each month one-third of the sample ED's are visited by the interviewers;
– in each month, one-sixth of the dwellings is replaced in the ED's pertainig to the month in consideration;
– any household entering the sample at some time is expected to provide labour market information on six consecutive occasions, then leaves the sample for ever;
– depending on earlier information on non-response rates, three or four addresses are selected from each ED visited in a given month;
– all individuals to be interviewed in the three or four dwellings within an ED visited currently have participated in the survey the same number of times (including non-responses), which means that they had entered the working sample at the same time, and will leave it together, too;
– in sampled dwellings, all persons aged 15–74 are eligible for the LFS.

The following remarks are in order here:

– while the above scheme is obviously complex, it guarantees that in principle there is an overlap of 5/6 between the samples of two consecutive quarters, an overlap of 2/3 between consecutive half-years, and 1/3 between same quarters of consecutive years;
– the complexity has resulted from exogeneous factors such as budget cuts and not from some extravagant new tendencies in sampling techniques;
– while it is possible to identify rotation groups in the system given, owing to the different number of dwellings selected from different ED's (three in some cases and four in others), there may be significant differences among rotation groups, indicating that composite estimators based heavily on them may be not very successful;
– a new redesign is underway to remove the asymetry among rotation groups.

By the sample design, the Horvitz–Thompson estimator is the natural tool to estimate levels or totals from the LFS. Since dwellings are the ultimate units of selection in the sample, primary sample weights are expressed as ratios of total number of dwellings in a given geographical unit to that in the sample. To reduce non-response bias, final value of sample weights is obtained by adjustment with the method of generalized iterative scaling.[3]

This method ensures that adjusted estimates for totals of age-sex groups and dwellings in geographical breakdown agree with the corresponding updated census counts. For further details on LFS design, see the following articles.[4]

The simple composite estimator and its stability

Composite estimators can be defined in different ways. Our interest is focused on such estimators which need for their definition sample data from at least two periods of time, say t and t+1, requiring also that the samples at those periods may have a non-empty overlap. If this is the case, the definition is as follows:

$$\hat{Y}^c_{t+1} = (1-\alpha_t)\hat{Y}_{t+1} + \alpha_t \hat{Y}^c_t + \alpha_t \Delta \hat{Y}'_t, \qquad /1/$$

where

$\hat{Y}_t$ – is the direct sample estimate of some total Y at period t,
$\hat{Y}^c_t$ – is the corresponding composite estimate,
$\hat{Y}'_t$ and $\hat{Y}'_{t+1}$ – are direct sample estimates at periods t and t+1, respectively, estimated on the overlap of the samples used at those periods,
$\Delta\hat{Y}'_t = \hat{Y}'_{t+1} - \hat{Y}'_t$,
$0 \le \alpha_t \le 1 \cdot$

With reference to the previous section, it is easily seen that the conditions of introducing this estimator in the Hungarian LFS are met; recall in particular that the overlap of the samples between two consecuitve periods, i.e. quarters amounts to approximately 5/6.

In the first applications of estimator /1/ in the American CPS , the weight α_t was chosen as 0.5 for each period, which was month in that case, and for all variables of interest. The latter were totals of employment (E), unemployment (UE) and civilian labour force (CLF) at national level as well as in certain breakdowns such as e.g. by race. Experience showed that that choice of α_t reduced the variance of the variables, without destroying their consistency; e. g. levels of (E) and (UE) totalled the level of (CLF), no matter if simple direct or composite estimates were considered. *Breau* and *Ernst*, studying the so-called generalized composite estimate, have determined the coefficients – that correspond to the weight α_t in the case of estimator /1/ – so that the variance of the composite estimate may be minimal. While to minimize variances is very attractive, and application of this principle to simple composite estimators is particularly simple, this approach involves considerable

[3] *Darroch, J. N. – Ratcliff, D.*: Generalized iterative scaling for log-linear models. Annals of Mathematical Statistics. 1972. Vol. 43. 1470–1480. p.; *Zaslavsky, A. M.*: Representing local area adjustment by reweighing of households. Survey Methodology. 1988. Vol. 14. 265–286. p.; *Zieschang, K. D.*: Sample weighing methods and estimation of totals in the consumer expenditure survey. *Journal of the American Statistical Association*. 1990. Vol. 85. 986–1001. p.

[4] *Éltető, Ö.*: The unified system of Household Surveys in Hungary. Paper presented at the seminar "International Comparison of Survey Methodologies". Athens. 30 March–1 April 1992.; *Mihályffy, L.*: The Unified System of Household Surveys in the decade 1992–2001. Statistics in Transition. 1994. Vol. 1. 443–462. p.

complexities. With this method, obviously different values are obtained for α_t if the variance of different characteristics, e.g. level of (E) and (UE) are minimized. Though individual composite estimates would be of minimal variance in this way, their consistency in terms of additivity would be lost. An alternative approach to this would be to optimize only for one variable, probably for level of (UE), and to use α_t determined in that way for all other variables (levels) involved. However, even this method may lead to undesirable consequences; this can be illustrated by a comment of some US Bureau of Labor Force officials, who claimed that they would refuse a set of estimates if the price of reducing variance of level of (UE) were accepting an increase in the variance of total (CLF).

In the history of composite estimates, little attention has been devoted to the issue of stability. By this we mean the following. Suppose we consider a not too long time horizon, e.g. eight consecutive quarters in the Hungarian LFS. In the first period, the composite estimate is set equal to the current direct sample estimate, and in periods 2–8 the definition /1/ is used. The weight α_t is determined by minimizing var$\left(\hat{Y}^c_{t+1}\right)$, the estimated variance of the composite estimate. The latter is a quadratic function of α_t, thus minimization is done by differentiating with respect to α_t, and setting the derivative equal to 0. We shall call the composite estimate stable over the given time horizon if the sequence

$$\alpha_1, \alpha_2, \ldots, \alpha_7$$

shows the pattern of the sum of a constant mean and a random disturbance. If stability were found for some variable, e.g. for level of (UE), the average of the α_t 's would probably yield a variance close to its minimum and could be used as a constant for some periods in the future. There is obviously no guarantee that this kind of stability, if once established, would last for ever; in any case, the method described in the following enables the user to monitor the behaviour of α_t over time.

The method of examining stability

Stability of the simple composite estimate will be examined below with Hungarian LFS data, beginning with the first quarter of 1995 and completing with the fourth quarter of 1996, thus a time horizon of eight consecutive periods will be considered. The reader may observe that the method can easily be extended to the case where the number T of periods involved is different from 8. Note that the lengthy technical derivations below serve as the documentation of our experimental computations, hence readers with less interest in such details may skip to the next section on results.

Because of the recurrent relation /1/, the simple composite estimator can be re-written as follows:

$$\hat{Y}^c_{t+1} = a_{t+1,1}\hat{Y}_1 + \ldots + a_{t+1,t+1}\hat{Y}_{t+1} + b_{t+1,1}\Delta\hat{Y}'_1 + b_{t+1,2}\Delta\hat{Y}'_2 + \ldots + b_{t+1,t}\Delta\hat{Y}'_t \qquad /2/$$

where $\hat{Y}^c_1 = \hat{Y}_1$, and

$$a_{t+1,1} = \prod_{\ell=1}^{t}\alpha_\ell, \quad a_{t+1,t+1} = (1-\alpha_t), \quad t = 1,2,\ldots,7 \qquad /3a/$$

$$a_{t+1,s} = (1-\alpha_{s-1})\prod_{\ell=s}^{t}\alpha_{\ell}, \quad t = 1,2,...,7, \quad s = 2,3,...,t \qquad /3b/$$

$$b_{t+1,s} = \prod_{\ell=s}^{t}\alpha_{\ell}, \quad t = 1,2,...,7, \quad s = 1,2,...,t \qquad /3c/$$

These relations can be verified by some algebra and induction. In what follows the above notations will be simplified, some additional notations will be introduced, and then an algorithm for the computation of the composite estimates in the periods 2, 3, ..., 8 will be given. Let

$$a_{11} = 1, \quad a_{12} = a_{13} = \ldots = a_{1,,15} = 0, \qquad /3d/$$

furthermore

$$\Delta\hat{Y}_t' = \hat{Y}_{t+8}, \quad t = 1,2,...,7,$$

$$a_{t+1,9} = b_{t+1,1}, \quad a_{t+1,10} = b_{t+1,2}, \quad \ldots, a_{t+1,t+8} = b_{t+1,t} \qquad /3e/$$

Finally, set a_{ts} = 0 for all (t, s) not occurring in relations /3a/–/3e/. Since $\hat{Y}_1^c = \hat{Y}_1$, with these notations we have

$$\hat{Y}_t^c = a_{t1}\hat{Y}_1 + ... + a_{tt}\hat{Y}_t + ... + a_{t,15}\hat{Y}_{15}, \qquad /4/$$

for t = 1, 2, ..., 8, or, in vector-matrix form,

$$\hat{Y}_{\bullet}^c = A\hat{Y}_{\bullet} \qquad /4'/$$

where the dot represents the subscript ranging from 1 to 8 in the case of the composite estimate and from 1 to 15 in the case of the variable $\hat{Y}_t$. A is the 8×15 matrix of the coefficients a_{ts}. If

$$C = (c_{ij}) = \begin{pmatrix} \mathrm{var}\hat{Y}_1 & \mathrm{cov}(\hat{Y}_1,\hat{Y}_2) & \ldots & \mathrm{cov}(\hat{Y}_1,\hat{Y}_{15}) \\ \mathrm{cov}(\hat{Y}_1,\hat{Y}_2) & \mathrm{var}\hat{Y}_2 & \ldots & \mathrm{cov}(\hat{Y}_2,\hat{Y}_{15}) \\ \vdots & & & \\ \mathrm{cov}(\hat{Y}_1,\hat{Y}_{15}) & \mathrm{cov}(\hat{Y}_2,\hat{Y}_{15}) & & \mathrm{var}\hat{Y}_{15} \end{pmatrix} \qquad /5/$$

is the covariance matrix of the variables $\hat{Y}_t$, /4/ or /4'/ implies the following useful relations for t = 1, 2, ..., 8:

$$\mathrm{cov}(\hat{Y}_t^c, \hat{Y}_s) = a_{t1}c_{1s} + ... + a_{tt}c_{ts} + ... + a_{t,15}c_{15,s} \qquad /6/$$

for s = 1, 2, ..., 15, and

$$\mathrm{var}\hat{Y}_t^c = \sum_{i=1}^{15}\sum_{j=1}^{15} c_{ij}a_{ti}a_{tj} \qquad /7/$$

On the basis of /3a/–/3e/, /5/, /6/ and /7/, the following algorithm can be given for computing composite estimates and their variance. Note that α_1 in Step 1 and α_t in Step 4 are determined by the requirement that the composite estimate may have minimal variance at periods 2 and t+1, respectively.

Algorithm

1. $\alpha_1 = \dfrac{c_{22} - c_{12} - c_{29}}{c_{22} + c_{11} + c_{99} - 2c_{12} - 2c_{29} + 2c_{19}}$

2. $t = 2,\quad a_{t1} = a_1,\quad a_{t2} = 1 - \alpha_1,\quad a_{t9} = \alpha_1$

3. Using /6/ and /7/, compute $\hat{Y}_t^c$, $\mathrm{cov}\left(\hat{Y}_t^c, \hat{Y}_{t+1}\right)$ and $\mathrm{cov}\left(\hat{Y}_t^c, \hat{Y}_{t+8}\right)$

4. $\alpha_t = \dfrac{c_{t+1,t+1} - \mathrm{cov}\left(\hat{Y}_t^c, \hat{Y}_{t+1}\right) - c_{t+1,t+8}}{c_{t+1,t+1} + \mathrm{var}\hat{Y}_t^c + c_{t+8,t+8} - 2\mathrm{cov}\left(\hat{Y}_t^c, \hat{Y}_{t+1}\right) - 2c_{t+1,t+8} + 2\mathrm{cov}\left(\hat{Y}_t^c, \hat{Y}_{t+8}\right)}$

5. Using /3a/–/3c/ and /3e/, compute the coefficients $a_{t+1,1}$, $a_{t+1,2}$, …, $a_{t+1,15}$. Set the coefficients $a_{t+1,s}$ not defined in this way equal to 0.

6. If $t < 7$, increase t by 1, and go back to step 3.

Results and some conclusions

Under the assumption that eight consecutive quarters of the LFS are considered, the following relation was derived in the preceding section for composite estimates of some level or total:

$$\hat{Y}_{\bullet}^c = A\hat{Y}_{\bullet} \qquad /4'/$$

where

$\hat{Y}_{\bullet}^c$ – is a column vector consisting of the composite estimates in the 8 quarters in consideration,

$\hat{Y}_{\bullet}$ – is a column vector consisting of direct sample estimates in the 8 quarters plus the 7 estimates of changes in level between consecutive quarters,

A – is a 8×15 coefficient matrix transforming $\hat{Y}_{\bullet}$ into $\hat{Y}_{\bullet}^c$.

The matrix A was determined by the requirement that variance of the composite estimate may be minimal in each quarter except for the first. This requirement implied that, except for the first row, each entry of A became a function of the covariance matrix C defined by /5/ of the variable of interest. It is worth while to point out that /4'/ does work even if A is not related to the variable Y of interest through this optimality criterion. For instance, if someone decides on minimizing the variance of the level of unemployment (UE) and on using the same weights in the composite estimates for the employed (E) as for the (UE), he/she may proceed as follows. First, using data of the (UE) and the algorithm of the preceding section, the matrix A and by /4'/ the composite estimates of (UE) are computed. Next, using /4'/ with A obtained as optimal for the (UE) and with data of the (E), composite estimates of (E) are determined. The latter will be consistent with composite estimates of (UE), but their variance may not be minimal; under unfavourable conditions, it can be even higher than that of the corresponding direct sample estimate. Eq. /7/ of the preceding section may be used to estimate the variance of composite estimates in this situation, too, however, special care should be taken of the fact that the coefficient matrix A comes from optimizing for the (UE). while the covariance matrix C belongs to the (E).

In the following, details and results of our actual computations are presented. The first set of computations related to quarterly estimates of level of unemployment, from the first quarter of 1995 to the fourth quarter of 1996.

Table 1 shows the covariance matrix C for the (UE). Because of the complexity of direct sample estimates in the LFS, variances and covariances were estimated by the stratified jackknife option of the VPLX software developed by Robert E. Fay at the US Bureau of the Census. The VPLX is a software product specifically designed for complex surveys in which variance estimation with traditional analytic methods is infeasible.

Table 1

Covariance matrix of the variables for unemployed (all entries scaled by 10^{-6})

Direct sample estimates*								Quarterly changes in level**						
$\hat{Y}_1$	$\hat{Y}_2$	$\hat{Y}_3$	$\hat{Y}_4$	$\hat{Y}_5$	$\hat{Y}_6$	$\hat{Y}_7$	$\hat{Y}_8$	$\Delta\hat{Y}_1'$	$\Delta\hat{Y}_2'$	$\Delta\hat{Y}_3'$	$\Delta\hat{Y}_4'$	$\Delta\hat{Y}_5'$	$\Delta\hat{Y}_6'$	$\Delta\hat{Y}_7'$
153.6	104.1	74.8	50.2	31.4	15.6	0.0	0.0	19.5	8.0	5.9	0.9	2.7	0.0	0.0
104.1	143.2	103.5	71.2	47.6	25.7	4.9	0.0	9.1	14.5	9.3	2.9	3.7	0.3	0.0
74.8	103.5	153.4	105.5	73.8	47.6	15.7	5.4	3.9	19.2	16.3	5.5	4.8	0.9	1.0
50.2	71.2	105.5	151.6	107.0	73.5	29.6	14.4	2.5	5.9	17.5	14.8	7.1	4.7	1.9
31.4	47.6	73.8	107.0	173.9	119.0	55.2	29.3	1.6	1.8	9.5	32.2	15.9	13.3	5.0
15.6	25.7	47.6	73.5	119.0	164.7	85.7	50.5	0.2	1.5	7.4	16.0	11.2	19.9	9.3
0.0	4.9	15.7	29.6	55.2	85.7	158.8	101.1	0.0	1.0	1.3	6.7	6.2	9.0	22.4
0.0	0.0	5.4	14.4	29.3	50.5	101.1	143.0	0.0	0.0	0.9	1.9	2.8	1.8	14.2
19.5	9.1	3.9	2.5	1.6	0.2	0.0	0.0	35.5	4.4	1.0	0.3	0.4	0.0	0.0
8.0	14.5	19.2	5.9	1.8	1.5	1.0	0.0	4.4	42.1	5.7	3.4	0.2	0.2	0.0
5.9	9.3	16.3	17.5	9.5	7.4	1.3	0.9	1.0	5.7	40.9	5.7	0.4	0.8	0.4
0.9	2.9	5.5	14.8	32.2	16.0	6.7	1.9	0.3	3.4	5.7	58.1	8.3	4.0	1.3
2.7	3.7	4.8	7.1	15.9	11.2	6.2	2.8	0.4	0.2	0.4	8.3	32.9	2.1	1.9
0.0	0.3	0.9	4.7	13.3	19.9	9.0	1.8	0.0	0.2	0.8	4.0	2.1	42.3	6.5
0.0	0.0	1.0	1.9	5.0	9.3	22.4	14.2	0.0	0.0	0.4	1.3	1.9	6.5	46.1

* Subscripts 1, 2, ..., 8 refer to the quarters 1/1995 to 4/1996, respectively.
** $\Delta\hat{Y}_1' = \hat{Y}_2' - \hat{Y}_1'$, subscripts 1 and 2 refer to the first and second quarters of 1995, etc.

It is important to have some evidence on reliability of covariance estimates, and this is accomplished in Table 2, which contains size of the samples used to estimate the covariances. It is useful to note here that the entry c_{ij} of C is estimated on a sample which is the overlap of the samples used to estimate the variables $\hat{Y}_i$ and $\hat{Y}_j$ repectively. Empty cells in Table 2 refer to cases where the corresponding samples have no overlap. Considering that the smallest entry in the table is 5.9 and this corresponds to 5900 persons, we may assume that on the basis of our samples reliable covariance estimates can be computed. We also note that determining the data sets representing the overlaps of samples used in covariance estimates needed considerable amount of computer time and tremendous space on the hard disk. Table 3 shows the coefficient matrix A transforming direct sample estimates into composite estimates; it is probably the structure which may deserve attention here.

The matrix consists of two blocks which are lower tiangular matrices, and the subdiagonal of the second block contains the weight α_t in the composite estimate in the consecutive quarters.

Table 2

Size of intersections of samples used in variance-covariance computations
(in thousands)

Quarters 1995				Quarters 1996				Quarters 1995			1995	Quarters 1996		
1	2	3	4	1	2	3	4	1–2	2–3	3–4	4–1	1–2	2–3	3–4
51.9	43.3	35.4	27.9	19.9	13.0			43.3	35.1	27.6	19.6	12.8		
43.3	**51.1**	42.5	34.8	26.6	19.5	6.0		43.3	42.6	34.5	26.1	19.3	6.0	
35.4	42.5	**50.4**	42.1	33.5	26.3	12.5	5.9	35.1	42.5	42.1	32.9	26.1	12.4	5.9
27.9	34.8	42.1	**49.9**	40.7	33.4	19.2	12.3	27.6	34.5	42.1	40.2	33.1	19.1	12.3
19.9	26.6	33.5	40.7	**49.7**	41.8	27.1	19.7	19.8	26.3	33.3	40.9	41.8	27.0	19.7
13.0	19.5	26.3	33.4	41.8	**49.9**	34.8	27.1	12.8	19.3	26.2	33.3	41.8	34.8	27.0
	6.0	12.5	19.2	27.1	34.8	**49.6**	41.4		5.9	12.4	19.2	27.0	34.8	41.5
		5.9	12.3	19.7	27.1	41.4	**49.5**			5.9	12.4	19.6	27.0	41.4
43.3	43.3	35.1	27.6	19.8	12.8			*43.3*	35.1	27.3	19.4	12.7		
35.1	42.6	42.5	34.5	26.3	19.3	5.9		35.1	*42.6*	34.5	25.9	19.1	5.9	
27.6	34.5	42.1	42.1	33.3	26.2	12.4	5.9	27.3	34.5	*42.1*	32.9	26.0	12.4	5.9
19.6	26.1	32.9	40.2	40.9	33.3	19.2	12.4	19.4	25.9	32.9	*40.9*	33.3	19.2	12.3
12.8	19.3	26.1	33.1	41.8	41.8	27.0	19.6	12.7	19.1	26.0	33.3	*41.8*	27.0	19.6
	6.0	12.4	19.1	27.0	34.8	34.8	27.0		5.9	12.4	19.2	27.0	*34.8*	27.0
		5.9	12.3	19.7	27.0	41.5	41.4			5.9	12.3	19.6	27.0	*41.5*

Table 3

The Matrix A
Transforming direct sample estimates and quarterly changes
pertainig to unemployed to composite estimates

Direct sample estimates*								Quarterly changes in level**						
$\hat{Y}_1$	$\hat{Y}_2$	$\hat{Y}_3$	$\hat{Y}_4$	$\hat{Y}_5$	$\hat{Y}_6$	$\hat{Y}_7$	$\hat{Y}_8$	$\Delta\hat{Y}_1'$	$\Delta\hat{Y}_2'$	$\Delta\hat{Y}_3'$	$\Delta\hat{Y}_4'$	$\Delta\hat{Y}_5'$	$\Delta\hat{Y}_6'$	$\Delta\hat{Y}_7'$
1.0000	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
.2073	.7927	.0	.0	.0	.0	.0	.0	.2073	.0	.0	.0	.0	.0	.0
.0593	.2268	.7139	.0	.0	.0	.0	.0	.0593	.2861	.0	.0	.0	.0	.0
.0161	.0614	.1933	.7292	.0	.0	.0	.0	.0161	.0775	.2708	.0	.0	.0	.0
.0050	.0190	.0598	.2254	.6908	.0	.0	.0	.0050	.0240	.0837	.3092	.0	.0	.0
.0015	.0057	.0181	.0681	.2088	.6978	.0	.0	.0015	.0072	.0253	.0934	.3022	.0	.0
.0005	.0019	.0061	.0230	.0706	.2360	.6618	.0	.0005	.0024	.0086	.0316	.1022	.3382	.0
.0001	.0005	.0017	.0063	.0194	.0647	.1814	.7259	.0001	.0007	.0023	.0087	.0280	.0927	.2741

Note. The rows correspond to composite estimates $\hat{Y}_1^c, \hat{Y}_2^c, \ldots, \hat{Y}_8^c$.

* Subscripts 1, 2, ..., 8 refer to the quarters 1/1995 to 4/1996, respectively.

** $\Delta\hat{Y}_1' = \hat{Y}_2' - \hat{Y}_1'$, subscripts 1 and 2 refer to the first and second quarters of 1995, etc.

The results – direct sample and composite estimates and standard errors for the level of unemployment in the quarters 1995–1996 – can be found in Table 4. While there is definite decrease in the value of standard error as the result of compositing, the extent of improvement is rather modest.

Table 4

Estimates of the level of unemployment from the Hungarian LFS

Period (Quarters)	Direct sample estimate		Composite estimate	
	Data	Standard Error	Data	Standard Error
		1995		
First	435 948	12 395	435 948	12 395
Second	413 304	11 966	413 398	11 703
Third	416 895	12 387	414 747	11 965
Fourth	410 960	12 313	409 432	11 890
		1996		
First	425 420	13 186	42 5981	12 669
Second	399 701	12 833	398 666	12 283
Third	404 828	12 602	397 639	11 572
Fourth	375 627	11 959	376 161	11 413

The behaviour of the weight α_t showed the following pattern over time:

2d Quarter 1995 $\longrightarrow$ 4th Quarter 1996

0.2073 0.2861 0.2708 0.3092 0.3022 0.3382 0.2741

thus it may be regarded as stable. Over a period of moderate length in the future, the use of $\alpha_t = 0.3$ may be recommended. This choice would probably result in a value close to the minimum of the variance in the case of composite estimate of the level of unemployment.

Table 5

Estimates of the level of employment from the Hungarian LFS

Period (Quarters)	Direct sample estimate		Composite estimate	
	Data	Standard Error	Data	Standard Error
		1995		
First	3 600 001	49 701	3 600 001	49 701
Second	3 625 227	50 930	3 628 048	48 028
Third	3 656 034	53 025	3 660 898	47 102
Fourth	3 675 000	53 517	3 672 254	46 279
		1996		
First	3 578 796	50 999	3 598 358	45 718
Second	3 615 967	51 315	3 623 733	45 009
Third	3 648 239	51 323	3 652 138	41 042
Fourth	3 702 475	51 735	3 698 408	41 360

Of the computations concerning level of employment, only the results are reproduced here, since there is nothing new in the properties of the matrices A and C in comparison with the previous case. It is remarkable that the variance-reducing effect of compositing is far more relevant here than in the case of level of unemployment. (See Table 5.) Unfortunately, the weight α_t seems to be far from being stable in this case, as is shown by the pattern of its varying over time:

2d Quarter 1995			$\longrightarrow$			4th Quarter 1996
0.4924	0.6200	0.6302	0.4442	0.6180	0.5523	0.7035

A possible reason for this phenomenon is some imperfection in the current practice of processing the data of the LFS. There is some evidence that stability would be found here too, if that imperfection were eliminated; hopefully, this improvement would not destroy the stability found in the previous case.